



1600

#30

RAW SEQUENCE LISTING

DATE: 10/23/2003

PATENT APPLICATION: US/09/125,031C

TIME: 11:21:37

Input Set : A:\0660-0139-0XPCT.ST25.txt

Output Set: N:\CRF4\10232003\I125031C.raw

3 <110> APPLICANT: LONGACRE-ANDRE, SHIRLEY
 4 ROTH, CHARLES
 5 NATO, FARIDABANO
 6 BARNWELL, JOHN
 7 MENDIS, KAMINI
 9 <120> TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF
 PLASMODIUM MSP-1
 11 <130> FILE REFERENCE: 0660-0139-0XPCT
 13 <140> CURRENT APPLICATION NUMBER: 09/125,031C
 14 <141> CURRENT FILING DATE: 1999-03-10
 16 <150> PRIOR APPLICATION NUMBER: PCT/FR97/00290
 17 <151> PRIOR FILING DATE: 1997-02-14
 19 <150> PRIOR APPLICATION NUMBER: FR96/01822
 20 <151> PRIOR FILING DATE: 1996-02-14
 22 <160> NUMBER OF SEQ ID NOS: 15
 24 <170> SOFTWARE: PatentIn version 3.1
 26 <210> SEQ ID NO: 1
 27 <211> LENGTH: 291
 28 <212> TYPE: DNA
 29 <213> ORGANISM: Artificial Sequence
 31 <220> FEATURE:
 32 <223> OTHER INFORMATION: SYNTHETIC
 34 <220> FEATURE:
 35 <221> NAME/KEY: CDS
 36 <222> LOCATION: (1)..(291)
 37 <223> OTHER INFORMATION:
 W--> 39 <400> 1
 40 gaa ttc aac atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag 48
 41 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
 42 1 5 10 15
 44 aac tct ggc tgt ttc aga cac ttg gac gag aga gag gag tgt aaa tgt 96
 45 Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
 46 20 25 30
 48 ctg ctg aac tac aaa cag gag ggc gac aag tgc gtg gag aac ccc aac 144
 49 Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
 50 35 40 45
 52 ccg acc tgt aac gag aac aac ggc ggc tgt gac gca gac gcc aaa tgc 192
 53 Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
 54 50 55 60
 56 acc gag gag gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt 240
 57 Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
 58 65 70 75 80
 60 acc aaa ccc gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc taa 288
 61 Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser

ENTERED

RAW SEQUENCE LISTING

DATE: 10/23/2003

PATENT APPLICATION: US/09/125,031C

TIME: 11:21:37

Input Set : A:\0660-0139-0XPCT.ST25.txt

Output Set: N:\CRF4\10232003\I125031C.raw

```

62                               85                               90                               95
64 taa                                                                    291
67 <210> SEQ ID NO: 2
68 <211> LENGTH: 95
69 <212> TYPE: PRT
70 <213> ORGANISM: Artificial Sequence
72 <220> FEATURE:
73 <223> OTHER INFORMATION: SYNTHETIC
75 <400> SEQUENCE: 2
77 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
78 1                               5                               10                               15
81 Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
82                               20                               25                               30
85 Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
86                               35                               40                               45
89 Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
90                               50                               55                               60
93 Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
94 65                               70                               75                               80
97 Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
98                               85                               90                               95
101 <210> SEQ ID NO: 3
102 <211> LENGTH: 279
103 <212> TYPE: DNA
104 <213> ORGANISM: Plasmodium falciparum
106 <400> SEQUENCE: 3
107 aacatttcac aacaccaatg cgtaaaaaaa caatgtccag aaaattctgg atgtttcaga      60
109 catttagatg aaagagaaga atgtaaattg ttattaaatt acaaacaaga aggtgataaa      120
111 tgtgttgaaa atccaaatcc tacttgtaac gaaaataatg gtggatgtga tgcagatgcc      180
113 aaatgtaccg aagaagattc aggtagcaac ggaaagaaaa tcacatgtga atgtactaaa      240
115 cctgattctt atccactttt cgatggtatt ttctgcagt      279
118 <210> SEQ ID NO: 4
119 <211> LENGTH: 354
120 <212> TYPE: DNA
121 <213> ORGANISM: Artificial Sequence
123 <220> FEATURE:
124 <223> OTHER INFORMATION: SYNTHETIC
126 <220> FEATURE:
127 <221> NAME/KEY: CDS
128 <222> LOCATION: (1)..(354)
129 <223> OTHER INFORMATION:
W--> 131 <400> 4
132 gaa ttc aac atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag      48
133 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
134 1                               5                               10                               15
136 aac tct ggc tgt ttc aga cac ttg gac gag aga gag gag tgt aaa tgt      96
137 Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
138                               20                               25                               30
140 ctg ctg aac tac aaa cag gag ggc gac aag tgc gtg gag aac ccc aac      144

```

RAW SEQUENCE LISTING

DATE: 10/23/2003

PATENT APPLICATION: US/09/125,031C

TIME: 11:21:37

Input Set : A:\0660-0139-0XPCT.ST25.txt

Output Set: N:\CRF4\10232003\I125031C.raw

```

141 Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
142          35          40          45
144 ccg acc tgt aac gag aac aac ggc ggc tgt gac gca gac gcc aaa tgc      192
145 Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
146      50          55          60
148 acc gag gag gag gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt      240
149 Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
150 65          70          75          80
152 acc aaa ccc gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc tcc      288
153 Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser Ser
154          85          90          95
156 tct aac ttc ttg ggc atc tcg ttc ttg ttg atc ctc atg ttg atc ttg      336
157 Ser Asn Phe Leu Gly Ile Ser Phe Leu Leu Ile Leu Met Leu Ile Leu
158          100          105          110
160 tac agc ttc att taa taa      354
161 Tyr Ser Phe Ile
162          115
165 <210> SEQ ID NO: 5
166 <211> LENGTH: 116
167 <212> TYPE: PRT
168 <213> ORGANISM: Artificial Sequence
170 <220> FEATURE:
171 <223> OTHER INFORMATION: SYNTHETIC
173 <400> SEQUENCE: 5
175 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
176 1          5          10          15
179 Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
180          20          25          30
183 Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
184          35          40          45
187 Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
188          50          55          60
191 Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
192 65          70          75          80
195 Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser Ser
196          85          90          95
199 Ser Asn Phe Leu Gly Ile Ser Phe Leu Leu Ile Leu Met Leu Ile Leu
200          100          105          110
203 Tyr Ser Phe Ile
204          115
207 <210> SEQ ID NO: 6
208 <211> LENGTH: 342
209 <212> TYPE: DNA
210 <213> ORGANISM: Plasmodium falciparum
212 <400> SEQUENCE: 6
213 aacatttcac aacaccaatg cgtaaaaaaa caatgtccag aaaattctgg atgtttcaga      60
215 catttagatg aaagagaaga atgtaaatgt ttattaaatt acaaacaaga aggtgataaa      120
217 tgtgttgaaa atccaaatcc tacttgtaac gaaaataatg gtggatgtga tgcagatgcc      180
219 aaatgtaccg aagaagattc aggtagcaac ggaaagaaaa tcacatgtga atgtactaaa      240

```

RAW SEQUENCE LISTING

DATE: 10/23/2003

PATENT APPLICATION: US/09/125,031C

TIME: 11:21:37

Input Set : A:\0660-0139-0XPCT.ST25.txt

Output Set: N:\CRF4\10232003\I125031C.raw

```

221 cctgattcctt atccactttt cgatgggtatt ttctgcagtt cctctaactt cttaggaata 300
223 tcattcttat taatactcat gttaatatta tacagtttca tt 342
226 <210> SEQ ID NO: 7
227 <211> LENGTH: 387
228 <212> TYPE: DNA
229 <213> ORGANISM: Plasmodium falciparum
231 <220> FEATURE:
232 <221> NAME/KEY: CDS
233 <222> LOCATION: (1)..(387)
234 <223> OTHER INFORMATION:
W--> 236 <400> 7
237 atg aag gcg cta ctc ttt ttg ttc tct ttc att ttt ttc gtt acc aaa 48
238 Met Lys Ala Leu Leu Phe Leu Phe Ser Phe Ile Phe Phe Val Thr Lys
239 1 5 10 15
241 gaa ttc aac atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag 96
242 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
243 20 25 30
245 gaa ttc aac atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag 144
246 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
247 35 40 45
249 aac tct ggc tgt ttc aga cac ttg gac gag aga gag gag tgt aaa tgt 192
250 Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
251 50 55 60
253 ctg ctg aac tac aaa cag gag ggc gac aag tgc gtg gag aac ccc aac 240
254 Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
255 65 70 75 80
257 ccg acc tgt aac gag aac aac ggc ggc tgt gac gca gac gcc aaa tgc 288
258 Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
259 85 90 95
261 acc gag gag gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt 336
262 Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
263 100 105 110
265 acc aaa ccc gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc taa 384
266 Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
267 115 120 125
269 taa 387
272 <210> SEQ ID NO: 8
273 <211> LENGTH: 127
274 <212> TYPE: PRT
275 <213> ORGANISM: Plasmodium falciparum
277 <400> SEQUENCE: 8
279 Met Lys Ala Leu Leu Phe Leu Phe Ser Phe Ile Phe Phe Val Thr Lys
280 1 5 10 15
283 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
284 20 25 30
287 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
288 35 40 45
291 Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
292 50 55 60

```

RAW SEQUENCE LISTING

DATE: 10/23/2003

PATENT APPLICATION: US/09/125,031C

TIME: 11:21:37

Input Set : A:\0660-0139-0XPCT.ST25.txt

Output Set: N:\CRF4\10232003\I125031C.raw

```

295 Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
296 65 70 75 80
299 Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
300 85 90 95
303 Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
304 100 105 110
307 Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
308 115 120 125

```

311 <210> SEQ ID NO: 9

312 <211> LENGTH: 330

313 <212> TYPE: DNA

314 <213> ORGANISM: Plasmodium falciparum

316 <220> FEATURE:

317 <221> NAME/KEY: CDS

318 <222> LOCATION: (1)..(330)

319 <223> OTHER INFORMATION:

W--> 321 <400> 9

```

322 gaa aca gaa agt tat aag cag ctt gta gcc aac gtg gac gaa ttc aac 48
323 Glu Thr Glu Ser Tyr Lys Gln Leu Val Ala Asn Val Asp Glu Phe Asn
324 1 5 10 15
326 atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag aac tct ggc 96
327 Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu Asn Ser Gly
328 20 25 30
330 tgt ttc aga cac ttg gac gag aga gag gag tgt aaa tgt ctg ctg aac 144
331 Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys Leu Leu Asn
332 35 40 45
334 tac aaa cag gag ggc gac aag tgc gtg gag aac ccc aac ccg acc tgt 192
335 Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn Pro Thr Cys
336 50 55 60
338 aac gag aac aac ggc ggc tgt gac gca gac gcc aaa tgc acc gag gag 240
339 Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys Thr Glu Glu
340 65 70 75 80
342 gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt acc aaa ccc 288
343 Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys Thr Lys Pro
344 85 90 95
346 gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc taa taa 330
347 Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
348 100 105

```

351 <210> SEQ ID NO: 10

352 <211> LENGTH: 108

353 <212> TYPE: PRT

354 <213> ORGANISM: Plasmodium falciparum

356 <400> SEQUENCE: 10

```

358 Glu Thr Glu Ser Tyr Lys Gln Leu Val Ala Asn Val Asp Glu Phe Asn
359 1 5 10 15
362 Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu Asn Ser Gly
363 20 25 30
366 Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys Leu Leu Asn
367 35 40 45

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/125,031C

DATE: 10/23/2003
TIME: 11:21:38

Input Set : A:\0660-0139-0XPCT.ST25.txt
Output Set: N:\CRF4\10232003\I125031C.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 9

VERIFICATION SUMMARY

DATE: 10/23/2003

PATENT APPLICATION: US/09/125,031C

TIME: 11:21:38

Input Set : A:\0660-0139-0XPCT.ST25.txt

Output Set: N:\CRF4\10232003\I125031C.raw

L:39 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:37
L:131 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:4,Line#:129
L:236 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:7,Line#:234
L:321 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:9,Line#:319